AFRL-RW-EG-TP-2011-011

Infrared Stationary Object Acquisition and Moving Object Tracking

Sengvieng Amphay
David Gray

Air Force Research Laboratory, Munitions Directorate
AFRL/RWGI
101 West Eglin Blvd
Eglin AFB, FL 32542-6810

March 2011

Conference Paper

AIR FORCE RESEARCH LABORATORY
MUNITIONS DIRECTORATE

■ **Air Force Materiel Command** ■ **United States Air Force** ■ **Eglin Air Force Base, FL 32542**

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
07-03-2011	Interim	Nov, 2008 – Sep, 2010

4. TITLE AND SUBTITLE

Infrared Stationary Object Acquisition and Moving Object Tracking

5a. CONTRACT NUMBER
NA

5b. GRANT NUMBER
NA

5c. PROGRAM ELEMENT NUMBER
61102F

6. AUTHOR(S)

Sengvieng Amphay
David Gray

5d. PROJECT NUMBER
2311

5e. TASK NUMBER
EW

5f. WORK UNIT NUMBER
91

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Air Force Research Laboratory, Munitions Directorate
AFRL/RWGI
101 West Eglin Boulevard
Eglin AFB, FL 32542-6810

8. PERFORMING ORGANIZATION REPORT NUMBER

AFRL-RW-EG-TP-2011-011

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Air Force Research Laboratory, Munitions Directorate
AFRL/RWGI
101 West Eglin Boulevard
Eglin AFB, FL 32542-6810

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL-RW-EG

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-RW-EG-TP-2011-011

12. DISTRIBUTION / AVAILABILITY STATEMENT

DISTRIBUTION A: Approved for public release; distribution unlimited. 96 ABW/PA Approval and Clearance # 96ABW-2010-0495, dated 7 September 2010.

13. SUPPLEMENTARY

SEE COVER PAGE FOR PERTINENT METADATA INFORMATION.

14. ABSTRACT

Currently, there is much interest in developing electro-optic and infrared stationary and moving object acquisition and tracking algorithms for Intelligence, Surveillance, and Reconnaissance (ISR) and other applications. Many of the existing EO/IR object acquisition and tracking techniques work well for good-quality images, when object parameters such as size are well-known. However, when dealing with noisy and distorted imagery many techniques are unable to acquire stationary objects nor acquire and track moving objects.

This technical paper will discuss two inter-related problems: (1) stationary object detection and segmentation and (2) moving object acquisition and tracking in a sequence of images that are acquired via an IR sensor mounted on both stationary and moving platforms.

15. SUBJECT TERMS

autonomous object acquisition, stationary object detection, moving object acquisition, object tracking, autonomous object recognition, moving object indicator

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Sengvieng Amphay
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL	12	**19b. TELEPHONE NUMBER** *(include area code)* 850-883-0883

Infrared Stationary Object Acquisition and Moving Object Tracking

Sengvieng Amphay and David Gray
Integrated Sensing and Processing Sciences Branch
U.S. Air Force Research Laboratory 101 West Eglin Blvd. Eglin AFB FL, USA 32542

ABSTRACT

Currently, there is much interest in developing electro-optic and infrared stationary and moving object acquisition and tracking algorithms for Intelligence, Surveillance, and Reconnaissance (ISR) and other applications. Many of the existing EO/IR object acquisition and tracking techniques work well for good-quality images, when object parameters such as size are well-known. However, when dealing with noisy and distorted imagery many techniques are unable to acquire stationary objects nor acquire and track moving objects.

This paper will discuss two inter-related problems: (1) stationary object detection and segmentation and (2) moving object acquisition and tracking in a sequence of images that are acquired via an IR sensor mounted on both stationary and moving platforms.

1. A stationary object detection and segmentation algorithm called "Weighted Adaptive Iterative Statistical Threshold (WAIST)" will be described. The WAIST algorithm takes any intensity image and separates object pixels from the background or clutter pixels. Two common image processing techniques are nearest neighbors clustering and statistical thresholding. The WAIST algorithm uses both techniques iteratively, making best use of both techniques. Statistical threshold takes advantage of the fact that object pixels will exist above a threshold based on the statistical properties of the known noise pixels in the image. The nearest neighbor technique takes advantage of the fact that when many neighboring pixels are known object pixels, the pixel in question is more likely to be a object pixel. The WAIST algorithm initializes the nearest neighbor parameters and statistical threshold parameters and adjusts them iteratively to converge to an optimal solution. Each iteration of the algorithm conservatively declares a pixel to be noise as the statistical threshold is raised. This algorithm has proven to segment objects of interest from noisy backgrounds and clutter. Results of the effort are presented.

2. For moving object detection and tracking we identify the challenges that the user faces in this problem; in particular, blind geo-registration of the acquired spatially-warped imagery and their calibration. For moving object acquisition and tracking we present an adaptive signal/image processing approach that utilizes multiple frames of the acquired imagery for geo-registration and sensor calibration. Our method utilizes a cost function to associate detected moving objects in adjacent frames and these results are used to identify the motion track of each moving object in the imaging scene. Results are presented using a ground-based panning IR camera.

Keywords: autonomous object acquisition, stationary object detection, moving object acquisition, object tracking, autonomous object recognition, moving object indicator

1.0 INTRODUCTION

This paper presents algorithms for detecting multiple stationary and moving objects via 2D WAIST algorithm for non-moving and adaptive change detection for moving objects and estimating/tracking their *individual* motion paths. The first part of this paper, the stationary object acquisition, begins with the principles governing 2D WAIST algorithm - iteratively applying both image processing techniques: nearest

neighbors clustering and statistical thresholding to optimize the detection algorithm robustness and performance. The second part of the paper, the moving object acquisition and tracking, uses 2D adaptive change detection in dual imagery as the basis of our MTI approach. Initial results with FLIR imagery on ground platforms are illustrated.

2.0 IR IMAGE PROCESSING AND STATIONARY OBJECT ACQUISITION ALGORITHM

2.1 Technical Description of WAIST algorithm

Recent technological advances in IR sensor manufacturing enable the fabrication of compact and high quality focal plane (FPA) array cooled/uncooled IR cameras, suitable as sensors for smart weapons and unmanned aerial vehicles (UAV) Intelligence, Surveillance, and Reconnaissance (ISR) applications. The technology is continuing to develop and is increasingly used by commercial and military sectors, therefore the cost of IR cameras continue to decrease for seeker applications.
In this research, a novel image processing and object acquisition algorithm for mobile stationary objects was investigated, developed, and tested against numerous 2-dimentional modality imaging sensor data sets. This object acquisition and segmentation algorithm, WAIST is illustrated in Figure 1 below:

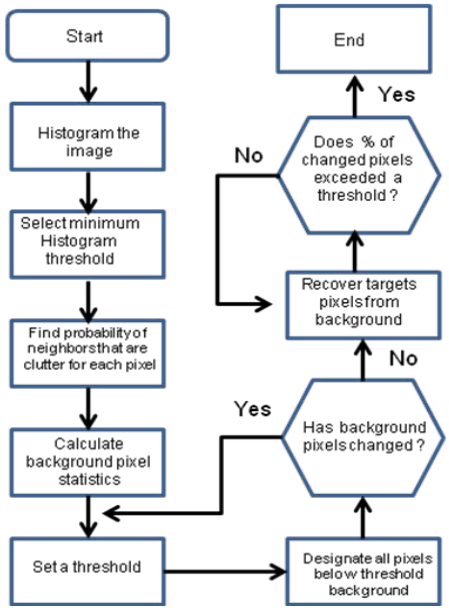

Figure 1 WAIST Algorithm

The WAIST algorithm begins with an initial assumption that at least a certain part of the image has objects of interest to be separated from the background, and that the objects of interest do not take up the entire scene. The algorithm designates an initial percentage of pixels in the image as background pixels. The algorithm then orders the pixels from lowest to highest intensity using the probability distribution function or histogram approach. A statistical adaptive threshold two parameter constant false alarm rate (CFAR) process is calculated to separate objects from background. This approach is developed under the assumption that the characteristics of the signal and noise change over different region of the image. In general, image characteristics differ considerably from one region to another. Degradations may also vary from one region to another. It is reasonable, then, to adapt the processing to the changing characteristics of the image and degradation. Therefore, the threshold, T_h initially assigns the pixels with low intensity signature to the background as described in the previous paragraph. During subsequent iterations, this

threshold is recomputed as a number of standard deviations times the fraction of nearest neighbor pixels that are object (not background) above the mean of the background pixels designated in the previous iteration. If the number of background pixels does not grow appreciably from one iteration to the next, than the algorithm is determined to have converged and the algorithm iterates no further.

Each iteration saves the resultant image to create a "data cube" that is a three dimensional array that is A by B by C, where A is the number of images produced and B by C is the size of the original image. Pixels that are deemed object pixels early in the algorithm can be assumed to be of a lower reliability than those that are deemed object pixels later in the algorithm. The number of iterations is automatically set by the algorithm itself based on the background statistical calculation. The threshold value is given by

$$T_h = \mu + W.m.\ \alpha \tag{1}$$

Were T_h = threshold

$$\mu = \frac{1}{N.M}\sum_1^N f(x_i)\sum_1^M f(x_j) \qquad i=1\ldots\ldots N, \text{ and } j = 1\ldots\ldots\ldots M \tag{2}$$

N = number rows of the image
M = number columns of the image
m = constant is depended on the quality of the image scene
W = Constant varying from 2 to 4 depended on the clutter probability distribution function

$$\alpha = \sqrt{\frac{\sum_{i-1}^{K}(x_i - x)^2}{K-1}} \tag{3}$$

K = sample size of the image

The fraction of nearest neighbor pixels in an n by n (n is another user set parameter) window surrounding but not including the pixel of interest is computed in another subroutine. It should be noted that the weighting of these nearest neighbor pixels is inversely proportional to their Euclidean distance from the center.

$$d_{i,j} = \sqrt{r_i^2 + c_j^2} \qquad I = 1\ldots\ldots..S, \text{ and } j = 1\ldots\ldots\ldots\ldots R \tag{4}$$

$$w_{i,j} = \frac{1}{d_{i,j}} \tag{5}$$

Where S is the size of row of kernel
R is the size of column of kernel
$d_{i,j}$ is the Euclidean distance from test pixel to neighbor pixel
$w_{i,j}$ is the weight of the neighbor pixel *i, j*

Note that when the threshold is established, all pixels below the threshold are deemed background pixels. This new set of background pixels is used to re-compute the mean and standard deviation for the next iteration. The new distribution of background pixels is also used to re-compute the nearest neighbor fraction of every image pixel. If the number of background pixels does not grow appreciably from one iteration to the next (this percentage is user set), then the algorithm is determined to have converged and the algorithm iterates no further.

Earlier, the WAIST algorithm was developed and exercised against several imaging sensor data sets. Recently, long-wave infrared data has also been evaluated using the WAIST object detection and segmentation algorithms. Figure 2 is a photo of stationary objects in a cluttered background. Figure 2 is its FLIR input image, which was used to test the WAIST algorithm. Figures 4a – 4c show that the WAIST

adaptive threshold was iteratively computed and how it separates the AOI from the background. As the number of iterations increase and a greater number of pixels are declared part of the background, the higher the confidence in the segmentation. Figure 5 illustrates the number of objects with labeling from 1 to 8 (8 objects acquired). Figure 6 shows the outcome of the segmentation. Furthermore, the algorithms also perform recursive iteration to further improve the final results of object segmentation by doing a check to ensure the algorithm didn't go too far and convert some useful pixels to background or "over converge". The final outcome of the algorithm is a segmentation of object from the background as shown in Figure 7. This output will help an analyst rapidly select the object from the background and it will increase the probability of object acquisition while decreasing the probability of missed object detection and PFA.

Figure 2 Scene picture

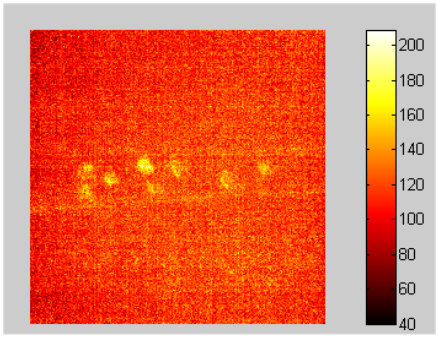

Figure 3 Original IR scene image

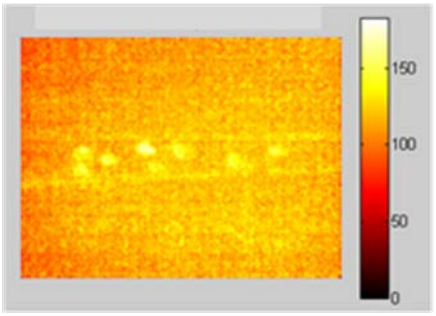

Figure 4a Result of 1st Iteration

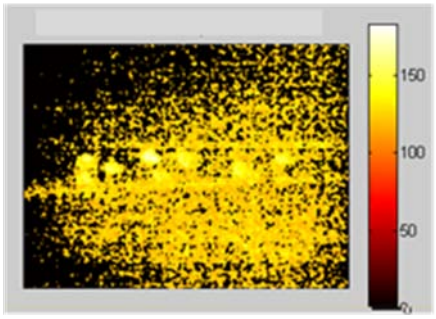

Figure 4b Result of 2nd iteration

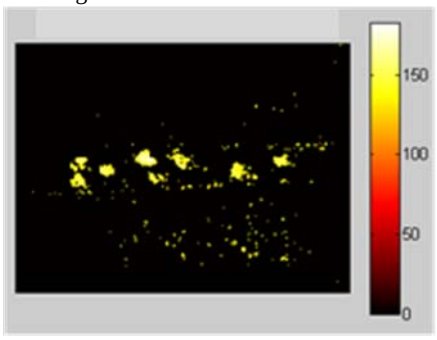

Figure 4c Result of 2nd iteration

Figure 5 Object labeling

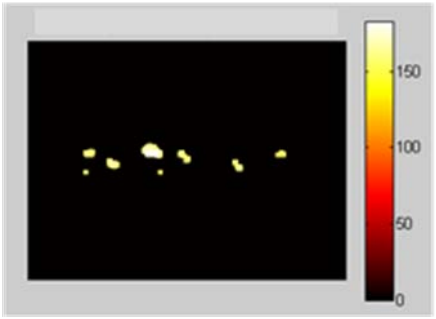

Figure 6 Object segmentation

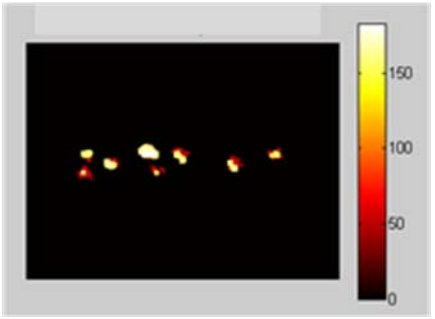

Figure 7 final Output from WAIST

3.0 APPLICATION OF 2D ADAPTIVE MTI IN FLIR IMAGERY

3.1 Signal Subspace Processing

A fundamental problem associated with these systems is that the *stationary background* should exhibit the same behavior (signature) when viewed by different sensory systems or at different time points. We refer to this scenario as perfectly *calibrated* sensors. Unfortunately, perfectly calibrated sensors do not exist in practice. Figure 8 represents a practical/realistic signal model for an uncalibrated dual sensory system that interrogates a scene that is composed of moving objects (change) as well as stationary objects.

In the ideal case of perfectly calibrated sensors, the change or MTI in two images can be detected by simply subtracting one image from the other. With uncalibrated sensors, the differencing operation is not practical. This is due to the fact that most of these dual sensory systems seek to detect subtle (weak) changes. Unfortunately, the power of the calibration error exceeds the power of a change in most practical scenarios.

Our approach for registering information in uncalibrated sensors is based on manipulating a system model with *unknown* parameters, which relates the outputs of two uncalibrated sensors, to develop a procedure to *blindly* calibrate the two outputs. This approach is based on a 2D adaptive filtering method that is identified in Figure 9. A practical method that does not require invention of large matrices, called Signal Subspace Processing (SSP), has been used to implement this 2D adaptive filter for radar platforms (Ref. [1]-[8]).

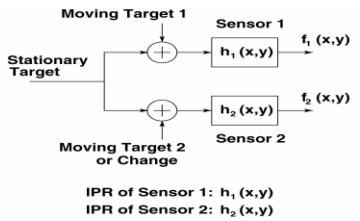

Figure 8 Signal model for dual sense imaging system

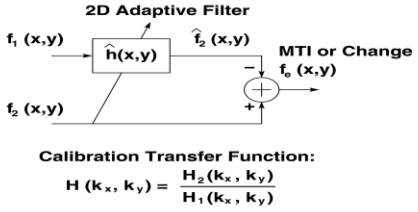

Figure 9 2D adaptive calibration of dual uncalibrated imagery

Between consecutive frames of an IR imaging sequence there are usually both camera motions and object motions. Before tracking moving objects, the effects of camera motions such as translation, rotation, zooming, panning, tiling and etc. need to be removed.

Our objective is to develop an MTI algorithm for time-series imagery from a visible or FLIR sensor on a stationary or moving platform. A change detection-based MTI algorithm that was originally developed for RF, *adaptively* (blindly) compensates for:

- Subtle rotation/scaling/shift (general spatial warping) of one image frame to another
- Camera (sensor) miscalibration and motion
- Subtle clutter (stationary objects) signature variations from one image frame to another

The basic signal model is identical to the one that we illustrate in Figure 8 for a RF platform. Sensor 1 is equivalent to a given frame; Sensor 2 corresponds to the image captured by another frame. If the camera does not move and is perfectly calibrated, then simple subtraction of the two channels (frames) is sufficient to construct the MTI. The blocks represented by $h_1(x, y)$ and $h_2(x, y)$ identify practical scenarios in which the sensor is moving during the data acquisition. This results in: a) viewing different stationary clutter background (gross shift between the two frames) that necessitates blindly identifying an appropriate *sweet spot* (that is, the scene that is common between the two frames) by the algorithm; b) *unequal* blur caused by the nonlinear motion of the camera that requires processing via 2D adaptive filtering.

It turns out that 2D versions of conventional adaptive filtering methods are computationally-intensive and/or require inversion of large matrices. A more practical algorithm, called Signal Subspace Processing (SSP), has been developed to address this issue. (This was originally developed for RF MTI and change detection Ref. [1]-[6].) We will examine this approach for an IR sensory system and provide results.

3.2 Camera motion estimation and stabilization

The basic hypothesis for this operation is that the relative coordinates of the camera for the present frame (e.g., Frame no. K that is called the *test* image) can be identified by *cross-correlating* this frame with a previous frame (e.g., Frame no. K-L1 that is called the *reference* image). Then, the relative motion between the test and reference images is estimated from the shift of the peak of their cross-correlation function from the origin. However, before performing the cross-correlation, two issues are addressed. First, the cross-correlation will possess a sharper/tighter peak in the spatial domain if the prominent feature of the two images is enhanced prior to cross-correlation. For this purpose, we use a high spatial frequency filter algorithm to enhance the edge of the test and reference images.

The second issue is related to the fact that the two frames that are cross-correlated do not record *identical* scenes due to the camera motion; that is, while most of the mid portions of the two images are the same, there are differences near the edges. Thus, if these different areas are not identified and removed prior to cross-correlation, the coordinates of the peak point might not be an accurate estimate of the relative shift between the two frames. Meanwhile, the common area (sweet spot) between Frames K and K-L1 is a function of the relative distance between the two frames that is unknown. However, the sweet spot between these two frames should be almost the same as the sweet spot between Frames K-1 and K-1-L1 (that is estimated earlier). Thus, we use this information to identify the sweet spots between Frames K and K-L1.

Once the relative shift between the current test Frame K and the reference Frame K-L1 is estimated, the overall shift of the current test frame from the first frame is found via adding this estimate to the estimate of the overall shift of Frame K-L1 that was estimated earlier. An important issue is clearly the choice of the lag that is used for the reference image, that is, L1. The key principle for this selection is that the common area (sweet spot) of the reference and test imagery should be sufficiently large for the cross-correlation processing to yield an accurate estimate of the relative camera shift. In this case, the selection of the lag parameter L1 depends on the relative speed of the camera motion with respect to the frame rate.

3.3 MTI via change detection

Once a current (test) frame is stabilized with respect to the previous frames, the next step is the generation of its Moving Object Indicator (MTI) image. For this purpose, the test image, Frame K, is compared to a previous Frame K-L2 (that is, a reference image) to detect *changes* in the current scene. For the change detection, the user may apply the 2D adaptive filtering method that we outlined earlier (Ref. [1]-[6]). In the case of IR and visible imagery that may contain warping, the adaptive filtering is essential. However, if the warping of visible or IR is nominal, a simple differencing is quite effective and computationally inexpensive. In this case, however, it might be useful to achieve a better spatial registration of the test and reference imagery via the cross-correlation method that we described in the previous section for the camera stabilization.

The rational for this is quite straightforward. In most cases, the camera motion also possesses slight rotation and scaling. The 2D adaptive filtering method does compensate for the subtle rotation and scaling of the test and reference imagery. However, when straight differencing is used, the slight rotation and scaling from one frame to another would result in a relatively small shift (for example, a couple of pixels) between Frames K and K-L2 even after the camera stabilization. (Recall that the camera stabilization Frame K is achieved via cross-correlating it with another reference image, that is, Frame K-L1.) Thus, prior to differencing, the cross-correlate of Frames K and K-L2 (that is, the test and reference images for change detection) are constructed to estimate their relative shift in the spatial domain. After compensating for this shift, the MTI is generated from the difference of the registered test and reference images.

Once an MTI image is created, the next step is to search this image for potential change or changes that represent moving objects. For this purpose, the peak of the MTI image is identified. If the value of the peak is greater than a pre-specified threshold, the algorithm decides that a moving object is present. A specific chip size around this peak is extracted. Using the moment method, the center of the gravity of the chip is determined, and recorded as the coordinates of a moving object in the test image. Then, the chip area around this moving object is nulled (zeroed) in the MTI image. The algorithm then loops back to the part where the peak of the MTI image is tested to determine if a moving object is present. After the peak value of the MTI image goes below the threshold, the task is completed with the coordinates of all moving objects recorded. The choice of the lag L2 is critical for the success of the change detection algorithm. The lag should be sufficiently large (in time) for a moving object to exhibit variations (changes) from Frame K-L2 to Frame K. Meanwhile, if the lag L2 is too large, then the common area (sweet spot) between the test and reference images becomes too small; that is, a large portion of the scene cannot be examined for MTI. As in the case of the camera stabilization, the frame rate plays an important role in the selection of the parameter L2. In the change detection problem, however, the relative speeds of the moving object and the frame rate should be considered to determine L2 (and not the speed of the camera motion since that is compensated for in the camera stabilization phase). As we mentioned before, the lag L2 should be sufficiently large such that a moving object in the imaging scene exhibits a shift due to a translational motion and/or any other motion (for example, waving arms) that is more than a couple of pixels from Frame K-L2 to Frame K; meanwhile, the lag L2 should be small enough to have a common area between Frames K-L2 and K that encompasses most (for example, 90 percent) of the imaging scene.

3.4 Results with an IR Camera

The camera stabilization and change detection algorithms are tested using a panning IR camera that is tracking two running individuals. Figures 10, 11, and 12 respectively, show the MTI, test image, and test image after being spatially-registered (stabilized) with respect to the first frame for Frames 2 and 78 of this database. The lag parameters that are used for this experiment are L1=1 and L2=5. The object chip size is 60 by 30 pixels.

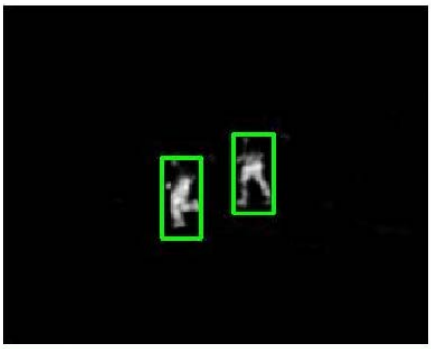

Figure 10 MTI image

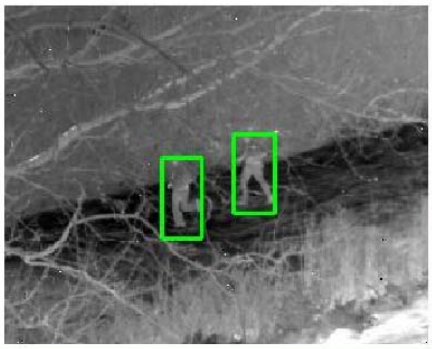

Figure 11 Test image

Figure 12 Panning FLIR camera: MTI, test image, and camera stabilized test image for frame 78

4.0 Multiple Moving Object Association and Tracking

4.1 Association and Tracking Algorithm

Before outlining the association and tracking algorithm, we first review the MTI algorithm for multiple moving objects. After an MTI image is constructed, the next step is to search this image for potential change or changes that represent moving objects. For this purpose, the peak of the MTI image is identified. If the value of the peak is greater than a pre-specified threshold, the algorithm decides that a moving object is present. The threshold is predefined based on statistical properties of the image scene, and then a specific chip size around this peak is extracted.

Using the moment method, the center of gravity of the chip is determined, and recorded as the coordinates of a moving object in the test image. Then, the chip area around this moving object is nulled (zeroed) in the MTI image. The algorithm then loops back to the part where the peak of the MTI image is tested to determine if a moving object is present. After the peak value of the MTI image goes below the threshold, the task is completed with the coordinates of all moving objects recorded; the outcome is illustrated in Figure 13-14 below.

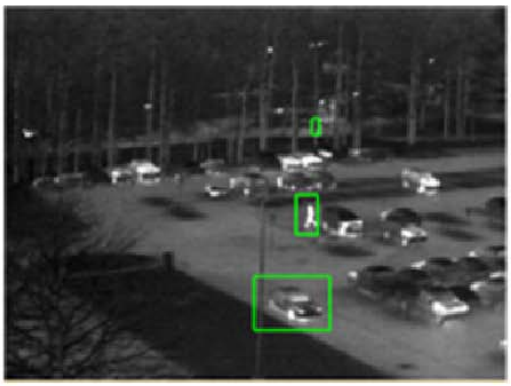

Figure 13 Multiple moving objects

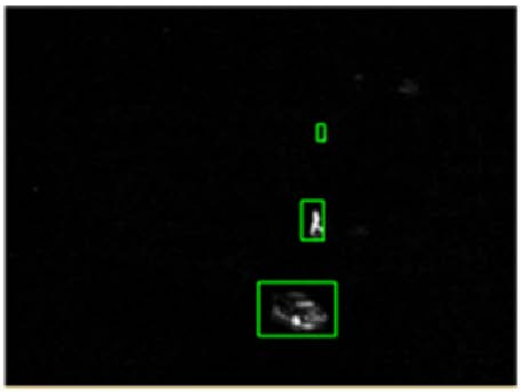

Figure 14 MTI of multiple moving objects

Note that the above-mentioned MTI algorithm detects multiple moving objects in the order of their *strength* in the MTI image. Thus, moving objects of the same type that have almost the same MTI signature levels are likely to have been detected in a *random* order. Thus, there is no information in the MTI image to clue the user to determine where a detected moving object in the present frame was in the previous frame and/or if a moving object just appeared in the scene.

The association and tracking algorithm resolves this ambiguity. The basic principles that are used to develop this algorithm are: a) there is *continuity* in the motion of a moving object (no random jumps) and b) a moving object does not move with a relatively high speed (with respect to the frame speed). Figure 15 exhibits how the algorithm works. This figure shows three associated object tracks in red, blue and green; these are shown with filled red, blue and green circles and smoothed dotted lines that approximately pass through them. Using linear prediction, the algorithm estimates (predicts) the coordinates of each of the three objects in the present (test) frame; these are shown by unfilled red, blue and green circles.

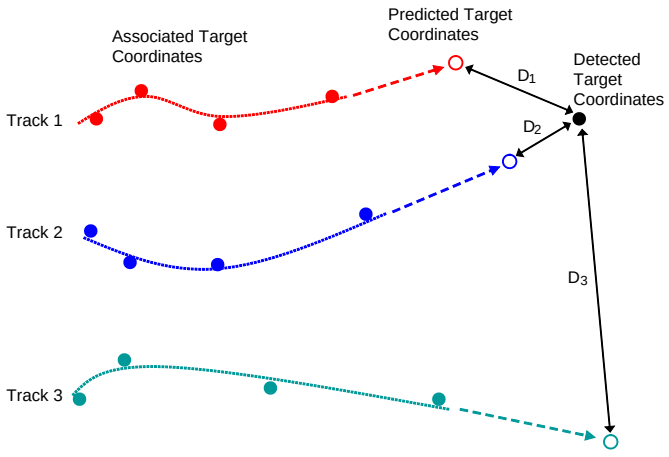

Figure 15 Depiction of Multiple Moving Object Track Association

Then, the algorithm identifies the distance between the coordinates of a detected moving object (that is shown by a black filled circle) in the test (current) frame to the predicted coordinates; these are identified as D1, D2 and D3. The algorithm selects the minimum distance that is D2 in the example in Figure 5. If this distance is less than a prescribed/pre-assigned maximum distance, call it Dmax, then the detected object (black filled circle) is associated with Track 2 (blue objects). However, if D2 is larger than Dmax, then the algorithm decides that the detected object is a new object in the scene, and starts a new object track for it. Dmax is set based on basic principle assumptions and constrains parameters of the object and the sensor as

described above. The algorithm tests all the detected objects in the present frame and associates them based on the above-mentioned minimum distance that is less than Dmax to the predicted object coordinates of the previously detected tracks. As we mentioned earlier, if a presently detected object cannot be associated with an existing track, a new object track is created. Meanwhile, if a predicted object cannot be associated with any of the detected objects in the present frame, then it is assumed to be *interrupted* but the track is not terminated. For those interrupted tracks, they are still treated as legitimate tracks. In fact, their predicted object coordinates are updated for the next frame and tested to be associated with the detected objects for the next frame. Once an interrupted object track is detected again, then a simple interpolation method (e.g., spline) is used to fill the gaps/interruptions in that track. Finally, all the tracks go through a smoothing/filtering operation to remove spikes and irregularities in a track, and yield a continuous-looking motion path for each moving object in the imaging scene.

4.2 Results with IR Camera

The multiple object association and tracking algorithm is tested using a panning IR camera Figure 12 that is tracking two running individuals. Figure 16 shows the associated tracks using the above-mentioned prediction and minimum distance algorithm. The output tracks are shown after interpolating and smoothing the interrupted tracks.

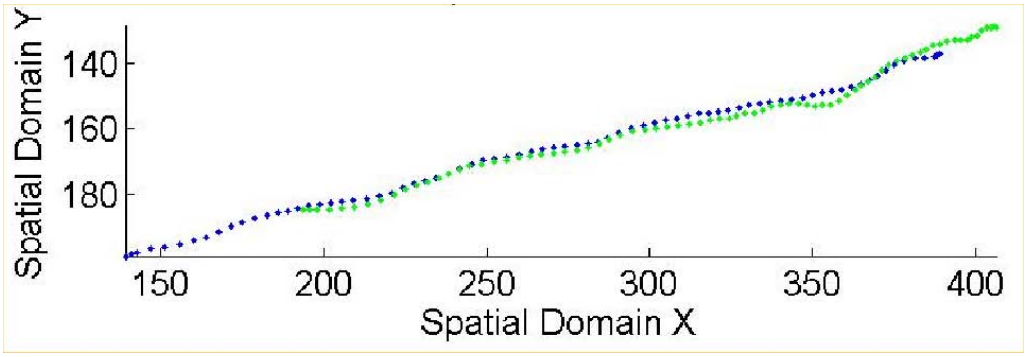

Figure 16 Filtered interpolated associated tracks of moving objects

5.0 GLOBAL SSP AND MULTI-FRAME SYNTHESIZED REFERENCE IMAGE

5.1 Analytical Foundation

The basic principle behind constructing change detection (CD) or MTI from two or multiple imagers that are acquired at different time points is that after compensating/calibrating for known (deterministic) differences of the two images and their spatial and spectral registration, the MTI image can be constructed via the following:

$$f_d(x,y) = f_T(x,y) - f_R(x,y), \tag{6}$$

where $f_T(x,y)$ and $f_R(x,y)$ are, respectively, the *test* image and the deterministically-calibrated *reference* image.

In practice, due to unknown variations of the camera electronics and platform coordinates, there are unknown image Impulse Response Function or Point Spread Function (IPR/PSF) variations and spatial warping in the acquired imagery that are unknown to the user. The simplest way to model this is to assume that these variations are invariant in the 2D domain of acquired imagery. In that case, under the null hypothesis, that is, there is no change or moving object, the test and reference images are related via the following:

$$\begin{aligned} \hat{f}_{RT}(x,y) &= f_R(x,y) \otimes h(x,y) \\ &= \int f_R(x-u, y-v)\; h(u,v)\, du\, dv \end{aligned} \tag{7}$$

where $\otimes$ represents two-dimensional convolution, and $h(x,y)$ is an unknown two-dimensional filter. Under the null hypothesis, this filter can be determined using the Linear Mean Square (LMS) algorithm; this approach is called *adaptive filtering*. A practical implementation of this method for the two-dimensional problems was described in our Ref. 2, and was referred to as Signal Subspace Processing (SSP).

The 2D complex image $\hat{f}_{RT}(x,y)$ is the LMS estimate of the test image from the reference image under the null hypothesis; we call $\hat{f}_{RT}(x,y)$ the *calibrated reference image*. The MTI is the constructed using the following:

$$f_d(x,y) = f_T(x,y) - \hat{f}_{RT}(x,y), \tag{8}$$

that yields zero under the null hypothesis (that is, no change or moving object). In presence of a change or a moving object, the LMS model is not valid. In this case, the estimate of the test image $\hat{f}_{RT}(x,y)$ is not equal to the test image $f_T(x,y)$. Hence, the difference of these two complex images yields a nonzero residual that signals the presence of a change or moving object.

A more realistic miscalibration model for the two receiver channels is based on the fact that the filter is *spatially-varying*. In this case, the relationship between the reference and test images can be expressed via the following:

$$\hat{f}_{RT}(x,y) = \int f_R(x-u, y-v)\ h_{xy}(u,v)\, du\, dv \tag{9}$$

where in this model the filter $h_{xy}(u,v)$ varies with the spatial coordinates, that is, (x,y). While the above model is a more suitable one, it is computationally prohibitive to implement the LMS or SSP method for this scenario.

A practical alternative is to assume that the filter is approximately spatially-invariant within a small area in the spatial domain. In this case, we can divide the test image into sub-patches within which the filter can be approximated to be spatially-invariant. The resultant model is

$$\begin{aligned} \hat{f}_{RT\ell}(x,y) &= f_{R\ell}(x,y) \otimes h_\ell(x,y) \\ &= \int f_{R\ell}(x-u, y-v)\ h_\ell(u,v)\, du\, dv \end{aligned} \tag{10}$$

where ℓ represent an index for the sub-patches.

In the approach that we call Local Signal Subspace Processing (LSSP), the LMS/SSP method is used to estimate the local unknown calibration filter $h_\ell(x,y)$. After this filter is estimated for each sub-patch, an approach that we call Global Signal Subspace Processing (GSSP) is used to estimate the original spatially-varying filter $h_{xy}(u,v)$ and the calibrated reference image (that is, estimate of the test image) via

. $$\hat{f}_{RT}(x,y) = \int f_R(x-u, y-v)\ h_{xy}(u,v)\, du\, dv \tag{11}$$

We have implemented a version of this algorithm, and are studying methods to improve its estimate.

In the case of multi-frame data, the user has access to multiple reference images, e.g., $f_R^{(n)}(x,y)$, $n = 1, \cdots, N$, where N is the number of reference images. Thus, for the n-th reference image, the calibrated reference image is

$$\hat{f}_{RT}^{(n)}(x,y) = \int f_R^{(n)}(x-u, y-v)\ h_{xy}^{(n)}(u,v)\, du\, dv, \tag{12}$$

$n = 1, \cdots, N$, where $h_{xy}^{(n)}(u,v)$ is the adaptive filter and is constructed via processing the test image $f_T(x,y)$ and the n-th reference image. $f_R^{(n)}(x,y)$

This procedure results in N calibrated reference images, $f_{RT}^{(n)}(x, y),\ n = 1, \cdots, N$. A critical issue in this investigation is how to combine these calibrated reference images to construct a common reference image, called *a Multi-Frame Synthesized Reference Image* that yields robust MTI imagery.

6.0 CONCLUSION

A good understanding of sensors specification and characteristic is very critical / essential for signal / image processing and algorithm development efforts, especially for IR camera. The qualities of IR imaging greatly influence the development, performance, and robustness of the signal / image processing algorithms. Image scene displays on computer monitor screen appear magnificent in the human observer point of view may not acceptable for machine vision. In some cases, the image scene that was enhanced contrast by employing the histogram equalization or other techniques has a poorer result than the raw data in testing of object detection algorithm. When the contrast of the object was enhanced from surrounding background, the intensity of pixels of clutter like objects were also proportionally increased, therefore it induced a high probability of false alarm (PFA). A novel adaptive/iterative image processing algorithm was developed exercised against mainly the raw data set, and its outcome of the WAIST algorithm appears very promising for fixed mobile object detection and segmentation.

This paper also described methods to detect moving objects in a sequence of imagery that was acquired via a visible or IR sensor on a moving platform. We presented an adaptive image processing method for blind geo-registration of the acquired spatially warped imagery and their calibration. We outlined a method to associate the detected moving objects in adjacent frames; the results were used to identify the motion track of each moving object in the imaging scene. Results were demonstrated on IR and visible imagery of ground and capture flight test data collection sets to exhibit the merits of the described methods.

7.0 REFERENCES

[1] Soumekh, M.,"Synthetic Aperture Radar Signal Processing with Matlab Algorithms, " New York: Wiley, 1999.

[2] Soumekh, M., "Signal Subspace Fusion of Uncalibrated Sensors with Application in SAR and Diagnostic Medicine," IEEE Transactions on Image Processing, vol. 8, no. 1, pp. 127-137, January 1999.

[3] Soumekh, M., "Moving Object Detection and Imaging Using an X Band Along-Track Monopulse SAR," IEEE Transactions on Aerospace and Electronic Systems, vol. 38, no. 1, pp. 315-333, January 2002.

[4] Soumekh, M. and Himed, B., "SAR-MTI Processing of Mutli-Channel Airborne Radar Measurement (MCARM) Data," Proceedings of IEEE International Conference on Radar Systems," pp. 24-28, Long Beach, May 2002.

[5] Dilsavor, R., Mitra, A., Hensel, M. and Soumekh, M., "GPS-Based Spatial and Spectral Registration of Delta Heading Multipass SAR Imagery for Coherent Change Detection," Proc. U.S. Army Workshop on Synthetic Aperture Radar Technology, Redstone Arsenal, October 2002.

[6] Soumekh, M., "Wavefront-Based Synthetic Aperture Radar Signal Processing," Frequenz, pp. 99-113, March/April 2001 special issue on SAR).

[7] Arredondo, F., Soumehk, M. and Amphay, S., "Sub-aperture and S-sub-band SAR Moving Object (MTI) Application," Proc. 55th Annual Meeting of the MSS Tri-Service Radar Symposium, Boulder, June 2009.

[8] Arredondo, F., Soumehk, M. and Amphay, S., "Integrated SAR Signal Processing for Imaging and Moving Object Indication," Proc. 54th Annual Meeting of the MSS Tri-Service Radar Symposium, Monterey, June 2008.

[9] Lim, J., "Two-Dimensional Signal and Image Processing," Prentice Hall, Inc 1990

[10] Lin, C., Amphay, S. and Sundstrom, B., "Sensor fusion with passive millimeter-wave and laser radar for object detection," Proceedings of SPIE, Vol. 3703, Jul 1999, pp 57 - 67.

[11] Gonzalez, R. and Woods, R, "Digital Image Processing" Pearson Prentice Hall, Inc. 2004

[12] Billard, D. and Brown C., "Computer Vision" Prentice Hall, Inc. 1982